Behind the criminal scenes

Criminal data analysis by

Andrea Koltai and Phil Busko

Crime Data Explorer

Source: https://crime-data-explorer.fr.cloud.gov/api

Common Fields

- Agency / State / National
- Year (2000-2018)
- Offense Type / Race / Gender
- Counts

Endpoints

- Police Employment (inc population)
- Arrests
- Offenses
- Offenders
- Victims

Employment Example

Agency	Year	Male_Officer	Male_Civilian

Arrests Example

State	Year	Gender

Offender Example

State	Year	Offense	Race

Crime Data Explorer

Is the police equally efficient in all US regions?

statistic of interest: $\frac{arrests}{population}$

$H_0 : \mu_{NE} = \mu_{MW} = \mu_{ST} = \mu_{WT}$

H_A : not all means are equal

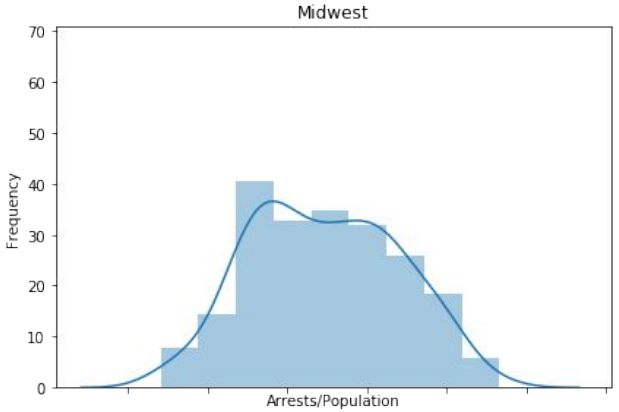
Method:

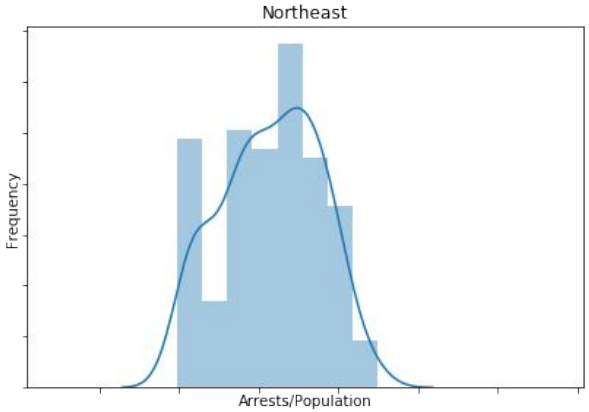
1. ANOVA
2. 2-sample t-test (x6)

	Region	States
0	Northeast	CT, MA, ME, NH, NJ, NY, PA, PR, RI, VT
1	Midwest	IA, IL, IN, KS, MI, MN, MO, ND, NE, OH, SD, WI
2	South	AL, AR, DC, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV
3	West	AK, AZ, CA, CO, HI, ID, MT, NM, NV, OR, UT, WA, WY

Crime Data Explorer

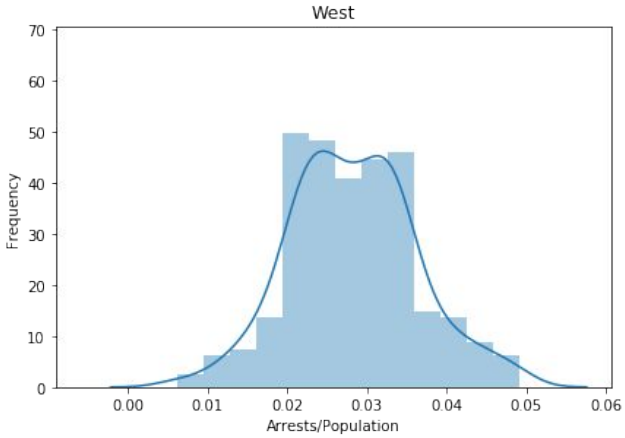
size = 220
ntest-p = 0.001

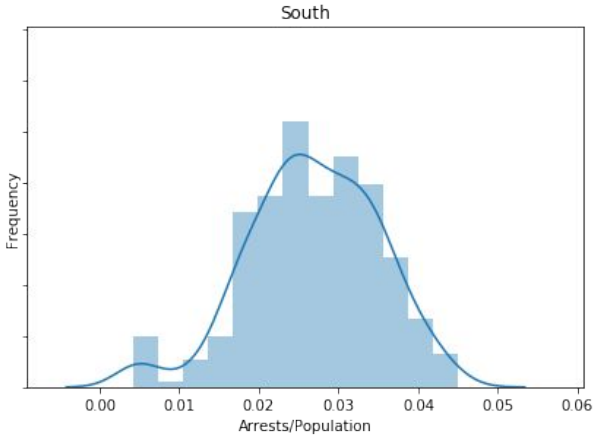




size = 171
ntest-p = 0.0001

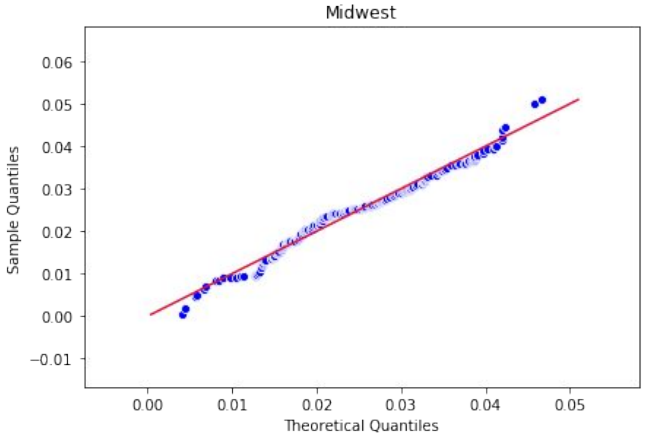
size = 244
ntest-p = 0.489

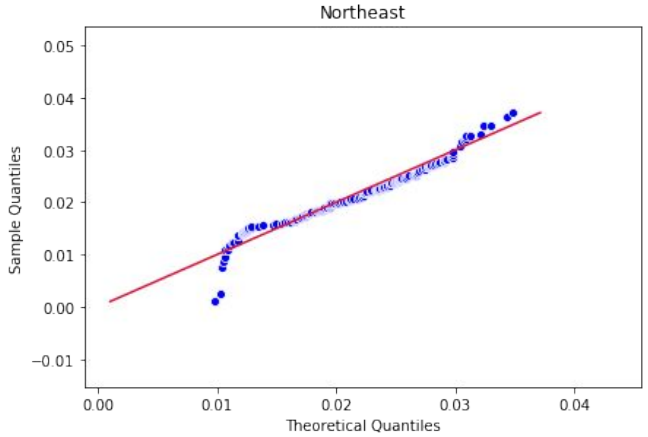


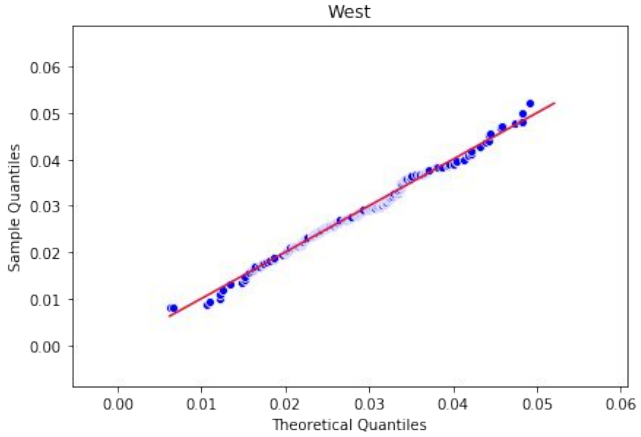


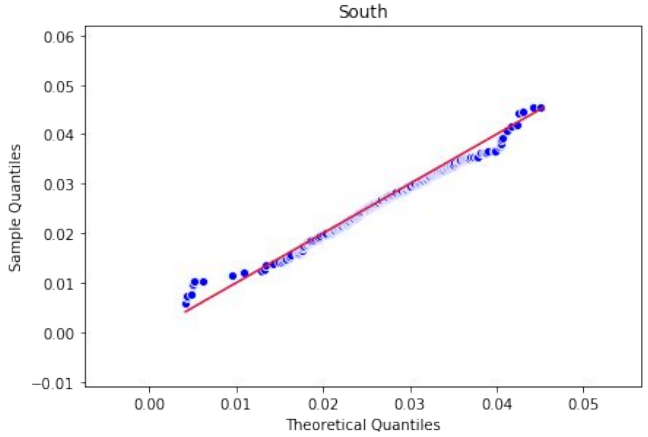
size = 288
ntest-p = 0.035

Crime Data Explorer









scipy.stats.f_oneway()

→ p = 9.67e-17

→ reject H_0

Crime Data Explorer

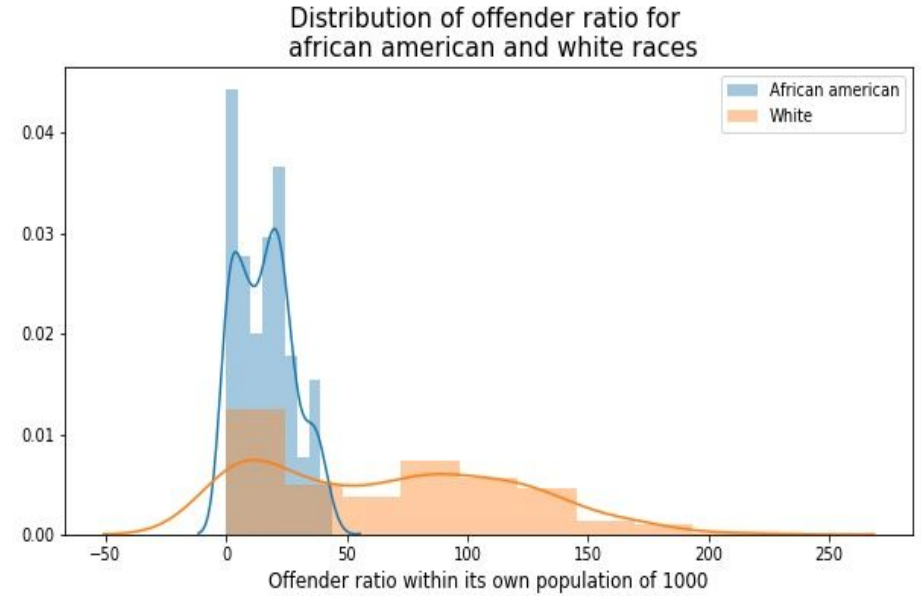
Next Question: Which region is the most underperforming ?

scipy.stats.ttest_ind()

```
midwest, northeast: 0.00012819
midwest, west: 7.75e-06
midwest, south: 0.0057939
northeast, west: 0.0
northeast, south: 0.0
west, south: 0.03622847
```

Crime Data Explorer

Question: are there differences in offenders ratio between African americans and white people among its population?



ratio calculation:
number of offenders / (population number x race ratio) x 1000

Average ratio for African americans: 67.48
Average ratio for white people: 16.20

Null hypothesis: there is no difference of mean of offender ratio between races
Alternative hypothesis: the ratio of offenders are higher among African American people
Significance level: 5%

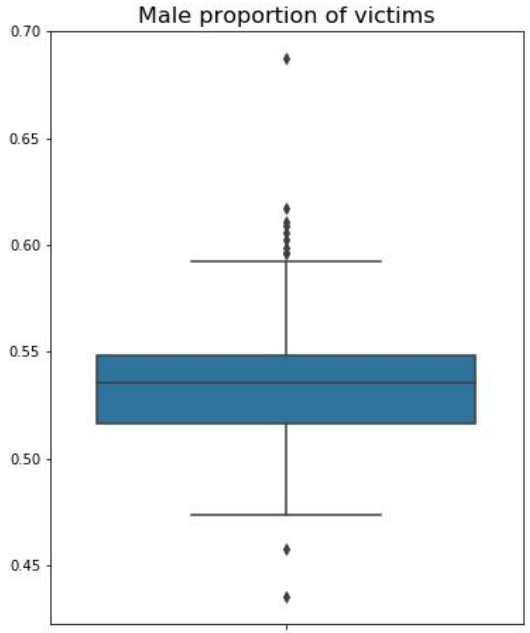
Hypothesis testing with paired T-test

z-score: 21.66 **p-value:** 0.00

Given the evidences there is significant difference in offender ratio between african american and white people.

Crime Data Explorer

Question: is there a difference in proportion of genders among victims?



Average proportion: 0.53
Standard deviation: 0.04

Null hypothesis: the proportion is 50%
Alternative hypothesis: the proportion is higher than 50%
Significance level: 5%

Hypothesis testing with Z-test for proportions

z-score: 136 **p-value:** 0.00

The proportion of male victims is significantly higher than 50%, the observed proportion is 53%.

Crime Data Explorer

Conclusions:

- By investigating the arrest data there was no significant difference in arrest ratio between regions.
- There is significant difference in crime rates between African American and White races among its own population.
- The proportion of male victims is significantly higher than 50%.

Next steps:

- Collecting individual crime data to understand the causes of bias between races.